DEPARTMENT OF STATISTICS

TECHNICAL REPORT No. 447

December, 1975

TOPICS IN TIME SERIES ANALYSIS

II. When are exponential smoothing forecast procedures optimal?

by

Johannes Ledolter

and

George E. P. Box

UNIVERSITY OF WISCONSIN-MADISON

Best
Available
Copy

When are exponential smoothing forecast procedures optimal?

2.1 Introduction and summary:

This chapter takes a critical look at exponential smoothing methods as described in section 1.6, particularly those of Brown which received broad attention in the literature and which have been widely used in practice. It will be shown that:

i) Brown's forecasting procedures are optimal in terms of achieving minimum mean squared error forecasts only if the underlying stochastic process is a specific member of the class of ARIMA(p,d,q) processes. Hence, it will be shown what assumptions are made when using these procedures.

ii) The implication of point (i) is that the stochastic processes which occur in the real world are from a specific subclass of ARIMA models. We, therefore, discuss the question of whether there are any reasons why these particular models should occur more frequently than others. There seems to be no satisfactory explanation.

iii) It is shown that even if the stochastic process which would lead to Brown's models occurred, the methods which he uses for making the forecasts are clumsy, and much simpler procedures can be obtained.

A related study has recently been reported independently by McKenzie (1973). The approach he takes is, however, different from the one adopted in this thesis.

2.2 Updating formulae for the coefficients in the forecast function:

Updating formula for Brown's forecasting procedures:

We saw in section 1.6 that Brown derives forecasts by specifying fitting functions (forecast functions) from the class of functions which satisfies

$$\underset{\sim}{f}(\ell + 1) = L\underset{\sim}{f}(\ell) \,. \qquad (2.2.1)$$

The transition matrix $L = \begin{bmatrix} L_{11} & \cdots & L_{1m} \\ \vdots & & \vdots \\ L_{m1} & \cdots & L_{mm} \end{bmatrix}$ is assumed to be nonsingular, and $\underset{\sim}{f}'(0) = [f_1(0),\dots,f_m(0)]$ is specified. The coefficients $\underset{\sim}{b}'(t) = [b_1(t),\dots,b_m(t)]$ of the forecast function

$$\hat{z}_t(\ell) = \hat{p}(t,\ell) = \sum_{i=1}^{m} b_i(t)f_i(\ell) = \underset{\sim}{b}'(t)\underset{\sim}{f}(\ell) \qquad (2.2.2)$$

are fitted by discounted least squares, minimizing

$$\sum_{j=0}^{t-1} \beta^j [z_{t-j} - \hat{p}(t,-j)]^2 .$$

In the <u>steady state</u> case $(t \to \infty)$, the solution is given by

$$\hat{\underset{\sim}{b}}(t) = F^{-1}\underset{\sim}{r}(t) \quad \text{where:} \quad F = \sum_{j=0}^{\infty} \beta^j \underset{\sim}{f}(-j)\underset{\sim}{f}'(-j)$$

$$\text{and} \quad \underset{\sim}{r}(t) = \sum_{j=0}^{\infty} \beta^j z_{t-j}\underset{\sim}{f}(-j). \qquad (2.2.3)$$

Brown (1962) developes an updating formula for the coefficients $\underset{\sim}{b}(t)$ from one time origin to the other:

$$\underset{\sim}{b}(t) = L'\underset{\sim}{b}(t-1) + \underset{\sim}{h}[z_t - \hat{p}(t-1,1)] \qquad (2.2.4)$$

where $z_t - \hat{p}(t-1,1)$ is the one step ahead forecast error,
and $\underset{\sim}{h} = F^{-1}\underset{\sim}{f}(0)$.

We note that $\underset{\sim}{h}$ is a function of the smoothing coefficient β only.

Updating formula for ARIMA (p,d,q) models:

The eventual forecast function of an ARIMA (p,d,q) process is the solution of $\phi_{p+d}(B)\hat{z}_t(\ell) = 0$ for $\ell > q$. It is given by

$$\hat{z}_t(\ell) = b_1^*(t)f_1^*(\ell) + \ldots + b_{p+d}^*(t)f_{p+d}^*(\ell) \quad \text{for } \ell > q-p-d. \qquad (2.2.5)$$

Box and Jenkins (1970) show that the updating formula for the coefficients $\underset{\sim}{b}^{*\prime}(t) = [b_1^*(t),\ldots,b_{p+d}^*(t)]$ is given by:

$$\underset{\sim}{b}^*(t) = L^{*\prime}\underset{\sim}{b}^*(t-1) + \underset{\sim}{g}[z_t - \hat{z}_{t-1}(1)] \qquad (2.2.6)$$

where: $L^{*\prime} = F_\ell^{*-1}F_{\ell+1}^*$

with
$$F_\ell^* = \begin{bmatrix} f_1^*(\ell) & \ldots & f_{p+d}^*(\ell) \\ f_1^*(\ell+p+d-1) & \ldots & f_{p+d}^*(\ell+p+d-1) \end{bmatrix}$$

and $\underset{\sim}{g} = F_\ell^{*-1}\underset{\sim}{\psi}_\ell$ with $\underset{\sim}{\psi}_\ell' = [\psi_\ell, \psi_{\ell+1}, \ldots, \psi_{\ell+p+d-1}]$
for any $\ell > q - p - d$.

ψ_i are the coefficients of $\psi(B) = \sum_{i=0}^{\infty} \psi_i B^i = \dfrac{\theta_q(B)}{\phi_{p+d}(B)}$.

Whenever $p + d = m$ and $f_i^*(\ell) = f_i(\ell)$ for $i = 1,2,\ldots,m$ it follows that

$$L^{*\prime} = F_\ell^{*-1}F_{\ell+1}^* = L' .$$

The matrix L appears, therefore, in the updating formula for the Brown model (2.2.4) as well as in the updating formula for the ARIMA model (2.2.6). This is because the transition matrix merely allows for changes in the coefficients arising from the change to a new origin, and has to occur in any reasonable updating formula. The essential difference between

the updating formulae (2.2.4) and (2.2.6), however, lies in the vectors $\underset{\sim}{h}$ and $\underset{\sim}{g}$. Whereas $\underset{\sim}{h}$ is a function of the smoothing parameter β only, $\underset{\sim}{g}$ is a function of all the parameters $(\underset{\sim}{\phi},\underset{\sim}{\theta})$ in the ARIMA model.

2.3 Equivalence theorems for Brown's forecast procedures and forecasts from ARIMA models:

We prove the following theorem.

Theorem 2.1:

Model A: (Brown model)

Consider the fitting functions

$$\underset{\sim}{f}'(\ell) = [u_1^{-\ell},\dots,u_n^{-\ell}] \text{ with } |u_i| \geq 1;\ u_i \neq u_j \text{ for } i \neq j;$$

$$\text{and } |u_i u_j| < 1/\beta \text{ for } 1 \leq i,\ j \leq n.$$

The coefficients of the forecast function

$$\hat{p}(t,\ell) = b_1(t)u_1^{-\ell}+b_2(t)u_2^{-\ell}+\dots+b_n(t)u_n^{-\ell} \qquad (2.3.1)$$

are estimated by discounted least squares with smoothing constant β; $0 < \beta < 1$.

Model B: ARIMA model

$$\prod_{i=1}^{n}\left(1 - \frac{1}{u_i}B\right)z_t = \prod_{i=1}^{n}(1 - u_i\beta B)a_t. \qquad (2.3.2)$$

Then: Brown's forecasting procedure using the fitting function as specified in model A will provide optimal fore-

casts in terms of minimizing the mean squared forecast error if the underlying stochastic process follows the ARIMA model given in (2.3.2).

Comment: This theorem will be proved by showing that model A and model B are equivalent in terms of having

i) the same <u>form</u> of the forecast function

ii) the same <u>updating</u> <u>formula</u> for the coefficients of the forecast function.

In order to show theorem 2.1 we will use the following lemma about the inverse of a Vandermonde matrix.

Lemma 2.1: Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_n \\ \vdots & & & \\ a_1^{n-1} & a_2^{n-1} & \dots & a_n^{n-1} \end{bmatrix} \text{where } a_i \neq a_j \text{ (for } i \neq j).$$

Then it is shown that the inverse of A is given by

$$A^{-1} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

where a_{ij} are the coefficients in the expansion

$$P_i(x) = \prod_{\substack{k=1\\k\neq i}}^{n} \frac{x-a_k}{a_i-a_k} = a_{i1}x^0 + a_{i2}x^1 + \dots + a_{in}x^{n-1} \quad (1\le i\le n) \tag{2.3.3}$$

The proof of Lemma 2.1 is given in the appendix to this chapter.

Proof of theorem 2.1:

a) The eventual forecast function for Model B is the solution of

$$\prod_{i=1}^{n}\left(1 - \frac{1}{u_i}B\right)\hat{z}_t(\ell) = 0 \quad \text{for} \quad \ell > n$$

and it is given by

$$\hat{z}_t(\ell) = b_1^*(t)u_1^{-\ell} + b_2^*(t)u_2^{-\ell} + \dots + b_n^*(t)u_n^{-\ell} \quad \text{for } \ell > 0. \tag{2.3.4}$$

The eventual forecast function (2.3.4) coincides with the forecast function of Model A given in (2.3.1).

b) The updating algorithm for the coefficients of the forecast function $\underset{\sim}{b}'(t)\underset{\sim}{f}(\ell)$ for Model A is given by

$$\underset{\sim}{b}'(t) = L'\underset{\sim}{b}(t-1) + \underset{\sim}{h}[z_t - \hat{p}(t-1,1)] \,. \tag{2.3.5}$$

Dobbie (1963) showed that for the case of exponential fitting

functions $\underset{\sim}{f}'(\ell) = [u_1^{-\ell},\dots,u_n^{-\ell}]$ where $u_i \neq u_j$ for $i \neq j$ and $|u_iu_j| < 1/\beta$ $(1\le i,j\le n)$ $\underset{\sim}{h}' = [h_1,\dots,h_n]$ is given by:

$$h_i = (1 - \beta u_i^2) \prod_{\substack{k=1\\k\neq i}}^{n} \frac{1-\beta u_i u_k}{1 - \frac{u_i}{u_k}} \,. \tag{2.3.6}$$

The updating algorithm for the coefficients of the eventual forecast function of the ARIMA model in (2.3.2) is given by:

$$\underset{\sim}{b}^*(t) = L'\underset{\sim}{b}^*(t-1) + \underset{\sim}{g}[z_t - \hat{z}_{t-1}(1)] \,. \tag{2.3.7}$$

Choosing $\ell = 1$ in (2.2.6), it is seen that

$$\underset{\sim}{g} = F_1^{*-1}\underset{\sim}{\psi}_1 = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & & & \\ a_1^n & a_2^n & \dots & a_n^n \end{bmatrix}^{-1} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{bmatrix} \tag{2.3.8}$$

where $a_i = \dfrac{1}{u_i}$ $\quad 1 \le i \le n$

and $\psi_k (1 \le k \le n)$ are the ψ-weights in

$$\psi(B) = 1 + \sum_{k=1}^{\infty} \psi_k B^k = \frac{\prod_{k=1}^{n} (1-\beta u_k B)}{\prod_{k=1}^{n} (1-\frac{1}{u_k} B)} \tag{2.3.9}$$

$$= \frac{\prod_{k=1}^{n} (1-\frac{\beta}{a_k} B)}{\prod_{k=1}^{n} (1-a_k B)} .$$

In order to prove the theorem we have to show that:

$\underset{\sim}{g} = \underset{\sim}{h}$ or equivalently

$$\Longleftrightarrow \begin{bmatrix} \frac{1}{a_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_2} & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & 0 & \frac{1}{a_n} \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix}^{-1} \underset{\sim}{\psi}_1 = \underset{\sim}{h}$$

$$\Longleftrightarrow \underset{\sim}{d}' \doteq \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ \vdots & & & \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix}^{-1} \underset{\sim}{\psi}_1 = \begin{bmatrix} a_1 h_1 \\ a_2 h_2 \\ \vdots \\ a_n h_n \end{bmatrix} \tag{2.3.10}$$

where $\underset{\sim}{d}' = [d_1, \ldots, d_n]$ and $d_i = a_{i1}\psi_1 + a_{i2}\psi_2 + a_{in}\psi_n$

and a_{ij} are the elements of the inverse of matrix A.

Using Lemma 2.1, we see that d_i is the coefficient of

x^0 in

$$\frac{1}{x}P_i(\frac{1}{x})\psi(x) = \frac{1}{x}[a_{i1} + a_{i2}\frac{1}{x} + \cdots + a_{in}\frac{1}{x^{n-1}}][1+\psi_1 x+\psi_2 x^2+\cdots] \tag{2.3.11}$$

$$= (a_{i1}\psi_1 + a_{i2}\psi_2 + \cdots + a_{in}\psi_n)x^0$$

$$+ \text{ terms with } x^{\ell} (\ell \neq 0).$$

Using the relation in (2.3.3) and (2.3.9)

$$\frac{1}{x}P_i(\frac{1}{x})\psi(x) = \frac{\frac{1}{x} \prod_{k \neq i} (\frac{1}{x} - a_k) \prod_{k=1}^{n} (1-\frac{\beta}{a_k} x)}{\prod_{k \neq i} (a_i - a_k) \prod_{k=1}^{n} (1-a_k x)}$$

$$= \frac{\frac{1}{x^n} \prod_{k \neq i} (1-a_k x) \prod_{k=1}^{n} (1-\frac{\beta}{a_k} x)}{\prod_{k \neq i} (a_i - a_k) \prod_{k=1}^{n} (1-a_k x)}$$

$$= \frac{1}{x^n(1-a_i x)} \frac{\prod_{k=1}^{n} (1-\frac{\beta}{a_k} x)}{\prod_{k \neq i} (a_i - a_k)} .$$

We, therefore, have to show that the coefficient of x^0 in

$$V_i(x) \doteq \frac{\prod_{k=1}^{n}(1-\frac{\beta}{a_k} x)}{x^n(1-a_i x)}$$

equals $\quad a_i \prod_{k\neq i} (a_i - a_k) h_i$. (2.3.12)

Using (2.3.6) we write

$$a_i \prod_{k\neq i} (a_i - a_k) h_i = a_i \left(1 - \beta \frac{1}{a_i^2}\right) \prod_{k\neq i} (a_i - a_k) \prod_{k\neq i} \frac{1 - \frac{\beta}{a_i a_k}}{1 - \frac{a_k}{a_i}}$$

$$= a_i^n \left(1 - \beta \frac{1}{a_i^2}\right) \prod_{k\neq i} (a_i - a_k) \prod_{k\neq i} \frac{1 - \frac{\beta}{a_i a_k}}{a_i - a_k}$$

$$= a_i^n \left(1 - \frac{\beta}{a_i^2}\right) \prod_{k\neq i} \left(1 - \frac{\beta}{a_i a_k}\right) .$$

It, therefore, remains to show that the coefficient of x^0 in $V_i(x)$ is equal to

$$a_i^n \left(1 - \frac{\beta}{a_i^2}\right) \prod_{k\neq i} \left(1 - \frac{\beta}{a_i a_k}\right) . \qquad (2.3.13)$$

Now

$$V_i(x) = \frac{\prod_{k=1}^{n} \left(\frac{1}{x} - \frac{\beta}{a_k}\right)}{(1 - a_i x)} = [1 + a_i x + a_i^2 x^2 + \ldots]$$

$$\times \left[\left(\frac{1}{x}\right)^n + c_1 \left(\frac{1}{x}\right)^{n-1} + \ldots + c_{n-1} \frac{1}{x} + c_n\right]$$

where the c_j's are the coefficients in the expansion of

$$\prod_{k=1}^{n} \left(\frac{1}{x} - \frac{\beta}{a_k}\right) ,$$

given by

$$\left\{\begin{array}{lcl} c_1 & = & -\beta \sum_{k=1}^{n} \frac{1}{a_k} \\ c_2 & = & \beta^2 \sum_{k<\ell} \frac{1}{a_k a_\ell} \\ c_3 & = & \beta^3 \sum_{k<\ell<m} \frac{1}{a_k a_\ell a_m} \\ \vdots & & \\ c_n & = & (-1)^n \frac{\beta^n}{a_1 a_2 \cdots a_n} \end{array}\right. . \qquad (2.3.14)$$

On the other hand, we expand

$$\prod_{k\neq i} \left(\frac{1}{x} - \frac{\beta}{a_i a_k}\right) = \left[\left(\frac{1}{x}\right)^{n-1} + e_1^{(i)} \left(\frac{1}{x}\right)^{n-2} + \ldots + e_{n-2}^{(i)} \frac{1}{x} + e_{n-1}^{(i)}\right]$$

where the coefficients $e_j^{(i)}$ $(1 \leq j \leq n-1)$ are given by

$$\left\{\begin{array}{lcl} e_1^{(i)} & = & -\frac{1}{a_i} \sum_{k\neq i} \frac{\beta}{a_k} \\ e_2^{(i)} & = & \frac{1}{a_i^2} \sum_{\substack{k<\ell \\ k\neq i\ \ell\neq i}} \frac{\beta^2}{a_k a_\ell} \\ e_3^{(i)} & = & -\frac{1}{a_i^3} \sum_{\substack{k<\ell<m \\ k\neq i, \ell\neq i, m\neq i}} \frac{\beta^3}{a_k a_\ell a_m} \\ \vdots & & \\ e_{n-1}^{(i)} & = & (-1)^{n-1} \frac{1}{a_i^{n-1}} \beta^{n-1} \prod_{k\neq i} \frac{1}{a_k} . \end{array}\right. \qquad (2.3.15)$$

It is readily shown that the relation between the coefficients in (2.3.14) and (2.3.15) is given by

$$\left\{\begin{array}{lll} c_1 & = & a_i e_1^{(i)} - \frac{\beta}{a_i} \\ c_2 & = & a_i^2 e_2^{(i)} - \beta e_1^{(i)} \\ c_3 & = & a_i^3 e_3^{(i)} - \beta a_i e_2^{(i)} \\ \vdots & & \\ c_{n-1} & = & a_i^{n-1} e_{n-1}^{(i)} - \beta a_i^{n-3} e_{n-2}^{(i)} \\ c_n & = & \qquad - \beta a_i^{n-2} e_{n-1}^{(i)} \end{array}\right. . \qquad (2.3.16)$$

Evaluating $\prod_{k\neq i} (\frac{1}{x} - \frac{\beta}{a_i a_k})$ at $x = 1$ gives

$$\prod_{k\neq i} (1- \frac{\beta}{a_i a_k}) = 1 + e_1^{(i)} + \ldots + e_{n-1}^{(i)}$$

and, therefore,

$$a_i^n (1- \frac{\beta}{a_i^2}) \prod_{k\neq i} (1- \frac{\beta}{a_i a_k}) = a_i^n (1- \frac{\beta}{a_i^2})(1+ e_1^{(i)} + \ldots + e_{n-1}^{(i)}) . \qquad (2.3.17)$$

Using the relation (2.3.16), the coefficient of x^0 in $V_i(x)$, let's say d_i^*, is given by

$$\begin{aligned} d_i^* &= a_i^n + a_i^{n-1} c_1 + \ldots + a_i c_{n-1} + c_n \\ &= a_i^n + a_i^{n-1} [a_i e_1^{(i)} - \frac{\beta}{a_i}] + a_i^{n-2} [a_i^2 e_2^{(i)} - \beta e_1^{(i)}] \\ &\quad + a_i^{n-3} [a_i^3 e_3^{(i)} - \beta a_i e_2^{(i)}] + \ldots \\ &\quad + a_i [a_i^{n-1} e_{n-1}^{(i)} - \beta a_i^{n-3} e_{n-2}^{(i)}] - \beta a_i^{n-2} e_{n-1}^{(i)} \\ &= a_i^n (1 - \frac{\beta}{a_i^2}) [1 + e_1^{(i)} + e_2^{(i)} + \ldots + e_{n-1}^{(i)}] \end{aligned} \qquad (2.3.18)$$

which equals (2.3.17), thus proving the claim that $\underset{\sim}{g} = \underset{\sim}{h}$.

<u>Theorem 2.2:</u>

Model A*: Brown forecasting procedure with fitting functions

$$\underset{\sim}{d}(\ell), \quad \text{where} \quad \underset{\sim}{f}(\ell) = R\underset{\sim}{d}(\ell). \qquad (2.3.19)$$

R is assumed to be nonsingular and $\underset{\sim}{f}'(\ell) = [u_1^{-\ell}, \ldots, u_n^{-\ell}]$ are the fitting functions of Model A in Theorem 2.1. The coefficients $\underset{\sim}{c}(t)$ of the forecast function $\underset{\sim}{c}'(t)\underset{\sim}{d}(\ell)$ are estimated by discounted least squares with smoothing coefficient $\beta (0 < \beta < 1)$.

Model B: ARIMA model

$$\prod_{i=1}^{n} (1- \frac{1}{u_i} B) z_t = \prod_{i=1}^{n} (1- \beta u_i B) a_t .$$

Then Brown's forecasting procedure using the fitting functions as specified in Model A* will provide optimal forecasts in terms of minimizing the mean squared forecast error if the underlying process follows the ARIMA Model B.

<u>Proof</u>:

For Model A*:

$$\underset{\sim}{b}'(t)\underset{\sim}{f}(\ell) = \underset{\sim}{b}'(t)R\underset{\sim}{d}(\ell) = \underset{\sim}{c}'(t)\underset{\sim}{d}(\ell) \quad \text{where } \underset{\sim}{c}(t) = R'\underset{\sim}{b}(t)$$

and

$$\underset{\sim}{b}(t) = L'\underset{\sim}{b}(t-1)+\underset{\sim}{h}[z_t-\hat{p}(t-1,1)]$$

$$R'\underset{\sim}{b}(t) = R'L'(R')^{-1}R'\underset{\sim}{b}(t-1)+R'\underset{\sim}{h}[z_t-\hat{p}(t-1,1)]$$

$$\underset{\sim}{c}(t) = R'L'R'^{-1}\underset{\sim}{c}(t-1)+R'\underset{\sim}{h}[z_t-\hat{p}(t-1,1)].$$

For Model B:

$$\underset{\sim}{b}^{*\prime}(t)\underset{\sim}{f}(\ell) = \underset{\sim}{b}^{*\prime}(t)R\underset{\sim}{d}(\ell) = \underset{\sim}{c}^{*\prime}(t)\underset{\sim}{d}(\ell)$$

where

$$\underset{\sim}{c}^*(t) = R'\underset{\sim}{b}^*(t)$$

and

$$\underset{\sim}{b}^*(t) = L'\underset{\sim}{b}^*(t-1) + \underset{\sim}{g}[z_t-\hat{z}_{t-1}(1)]$$

$$R'\underset{\sim}{b}^*(t) = R'L'R'^{-1}R'\underset{\sim}{b}^*(t-1)+R'\underset{\sim}{g}[z_t-\hat{z}_{t-1}(1)]$$

$$\underset{\sim}{c}^*(t) = R'L'R'^{-1}\underset{\sim}{c}^*(t-1)+R'\underset{\sim}{g}[z_t-\hat{z}_{t-1}(1)]$$

and since $\underset{\sim}{g} = \underset{\sim}{h}$ (from Theorem 2.1) it follows that $R'\underset{\sim}{g} = R'\underset{\sim}{h}$.

q.e.d.

Theorem 2.2 is important since sinusoidal fitting functions, which are frequently considered by Brown, can be written as linear combinations of exponential functions. For example,

$$\hat{p}(t,\ell) = c_1(t)+c_2(t)\sin\omega\ell+c_3(t)\cos\omega\ell$$

can be written as

$$\underset{\sim}{f}(\ell) = R\underset{\sim}{d}(\ell) \qquad (2.3.20)$$

where

$$\begin{bmatrix} 1 \\ e^{-i\omega\ell} \\ e^{i\omega\ell} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -i & 1 \\ 0 & i & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \sin\omega\ell \\ \cos\omega\ell \end{bmatrix} .$$

Equation (2.3.20) shows that the roots of the characteristic equation for the ARIMA model which is implied by the sinusoidal fitting functions in Brown's forecasting scheme are lying on the unit circle. For the case of monthly observations with $\omega = \frac{2\pi}{12}$, the autoregressive operator of the corresponding ARIMA model is given by $(1 - \sqrt{3}B + B^2)(1-B)$.

Many nonstationary time series exhibit homogeneity in the sense that apart from local level and/or trend and/or periodicity one part of the series behaves very much like

the other. Stochastic processes which exhibit these characteristics have some roots of $\phi(B) = 0$ on the unit circle. This implies that the autoregressive operator will contain factors of the form $(1-B), (1-B^2), (1-\sqrt{3}B+B^2), (1-B^s)$, etc., which Box and Jenkins refer to as simplifying operators.

It was shown in Table 1.1 that the $(1-B^{12})$ operator removes the effect of 12 equispaced roots on the unit circle. The eventual forecast function of the ARIMA model

$$(1 - B^{12})z_t = \theta_q(B)a_t$$

has 12 coefficients which are updated as each new observation becomes available.

Box and Jenkins point out that for monthly data with strong sinusoidal trend (e.g.: temperature data) an appropriate simplifying operator might be $(1 - \sqrt{3}B + B^2)$. For general 12-monthly seasonal patterns a complete set of sinusoidals can be achieved by considering the simplifying operator $(1 - B^{12})$.

Box and Jenkins also stress the fact that overdifferencing ought to be avoided, since it leads to non parsimonious, non invertible models. As an example we consider the invertible model

$$(1-B)(1-\sqrt{3}B+B^2)z_t = \theta_q(B)a_t. \tag{2.3.22}$$

The eventual forecast function of the model in (2.3.22) is a simple 12 point sinusoidal function

$$\hat{z}_t(\ell) = b_1^*(t)+b_2^*(t)\sin\frac{2\pi}{12}\ell + b_3^*(t)\cos\frac{2\pi}{12}\ell .$$

Overdifferencing the series by introducing the incorrect simplifying operator $(1-B^{12})$ results in a non-invertible model

$$(1-B^{12})z_t = \theta_q(B)(1-B+B^2)(1+B^2)(1+B+B^2)(1+\sqrt{3}B+B^2)(1+B)a_t . \tag{2.3.23}$$

A further discussion of this point is given by Abraham (1975).

In the following corollary to Theorem 2.1 and Theorem 2.2 we use the fact that, if the characteristic equation has a complex root, the conjugate complex will be a solution too.

<u>Corollary 2.1</u>

Model B: We consider the ARIMA model

$$\phi_n(B)z_t = \phi_n(\beta B)a_t \tag{2.3.24}$$

where the coefficients of $\phi_n(B)$ are real. Furthermore,

it is assumed that the roots of $\phi_n(B) = 0$ are distinct and lie on the unit circle, and that the eventual forecast function is given by

$$\hat{z}_t(\ell) = b_1^*(t)f_1(\ell)+\ldots+b_n^*(t)f_n(\ell).$$

Model A: Consider the Brown forecasting procedure with fitting functions

$$\underset{\sim}{f}'(\ell) = [f_1(\ell),\ldots,f_n(\ell)].$$

The coefficients of the forecast function

$$\hat{z}_t(\ell) = \underset{\sim}{b}'(t)\underset{\sim}{f}(\ell)$$

are fitted by discounted least squares with smoothing coefficient β, $0 < \beta < 1$.

Then the Brown forecasting procedure with fitting functions as specified in Model A will provide minimum mean squared error forecasts if the underlying stochastic process is given by the ARIMA Model B.

Theorem 2.3 will relax the condition of distinct roots.

<u>Theorem 2.3</u>:

Model B:

$$\phi(B)z_t = \phi(\beta B)a_t$$

where the coefficients of $\phi(B)$ are real and where the

roots of $\phi(B) = 0$ lie on the unit circle; furthermore, we allow the possibility of multiple roots.

The evenutal forecast function is given by

$$\hat{z}_t(\ell) = b_1^*(t)f_1(\ell)+\ldots+b_n^*(t)f_n(\ell).$$

Model A:

Brown model with fitting functions

$$\underset{\sim}{f}'(\ell) = [f_1(\ell),\ldots,f_n(\ell)]$$

and the coefficients of the forecast function are estimated by discounted least squares with smoothing parameter β.

Then Model A and Model B are equivalent in terms of having

a) the same form of forecast function

b) the same recursive updating formula for the coefficients of the forecast function $\underset{\sim}{b}'(t)$.

<u>Proof</u>: <u>ada</u>: Trivially the form of the forecast function of Model A and Model B is the same.

<u>adb</u>: We can write

$$\phi(B) = \prod_{i=1}^{s} \phi_i(B) \quad \text{where} \quad \phi_i(B) \quad 1 \leq i \leq s$$

are real valued polynomials in B with distinct roots on the unit circle;

Model B can be written as

$$\phi_1(B)\,\phi_2(B)\ldots\phi_s(B)z_t = \phi_1(\beta B)\,\phi_2(\beta B)\ldots\phi_s(\beta B)a_t$$

or

$$a_t = \frac{\phi_1(B)}{\phi_1(\beta B)}\,\frac{\phi_2(B)}{\phi_2(\beta B)}\cdots\frac{\phi_s(B)}{\phi_s(\beta B)}z_t.$$

We define $\quad a_t^{(s)} = \frac{\phi_s(B)}{\phi_s(\beta B)}\,z_t$, and through continued application of Corollary 2.1, the claim is proved.

q.e.d.

2.4 Interpretation of the results

Summary of the equivalence theorems:

Brown's forecasting procedures with specified fitting functions and smoothing parameter β will give minimum mean squared error forecasts if the underlying process follows the ARIMA model

$$\phi(B)z_t = \phi(\beta B)a_t. \qquad (2.4.1)$$

The roots of $\phi(B) = 0$ lie on the unit circle and the eventual forecast function of (2.4.1) is given by the fitting functions of Brown's model.

Shortcomings of Brown's exponential smoothing methods:

In light of the above theorem the shortcomings of Brown's forecasting procedure are threefold and they are summarized below:

i) The fitting functions of Brown's method which determine the form of $\phi(B)$ are picked by unreliable identification procedures. The form of the auto-regressive operator is in fact decided by the choice of the fitting functions and cannot be safely chosen by visual inspection of the time series itself. More reliable identification tools such as sample autocorrelation and sample partial autocorrelation function have to be considered.

A quadratic fitting function might well be used, and is actually used by Brown, to fit short pieces of data generated by a random walk model. If such fitting were relevant to forecasting, one might conclude that a polynomial of second degree was indicated. The random walk model, however, would actually lead to a polynomial forecast function of degree zero.

ii) The exponentially discounted weighted least squares procedure forces the moving average

operator, the right hand side of model (2.4.1), to be of the form $\phi(\beta B)$. It is thus automatically determined by the autoregressive part on the left hand side of model (2.4.1) and is a function of the smoothing constant only.

iii) The smoothing constant β is assumed to be known. Brown states that the smoothing constant should be picked between .7 and .9. Actual study of time series, however, gives no empirical support to this assertion and no theoretical reasons seem to be available for discussion. The supposition that β ought to be picked in this range appears strange.

The π-weights implied by Brown's model:

The π-weights for the ARIMA model (2.4.1) are derived by equating coefficients in

$$\frac{1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_n B^n}{1 - \phi_1 \beta B - \phi_2 \beta^2 B^2 - \dots - \phi_n \beta^n B^n} = 1 - \sum_{j=1}^{\infty} \pi_j B^j. \quad (2.4.2)$$

The weights are given by:

$$\begin{aligned} \pi_1 &= (1-\beta)\phi_1 \\ \pi_j &= \beta\phi_1\pi_{j-1} + \beta^2\phi_2\pi_{j-2} + \dots + \beta^{j-1}\phi_{j-1}\pi_1 + (1-\beta^j)\phi_j && 2 \le j \le n \\ \pi_j &= \beta\phi_1\pi_{j-1} + \beta^2\phi_2\pi_{j-2} + \dots + \beta^n\phi_n\pi_{j-n} && j \ge n+1 \end{aligned} \quad (2.4.3)$$

It is instructive to look at the π-weights since they show how past observations are discounted to derive minimum mean squared error forecasts. The π-weights will be plotted for a number of examples considered in the next section.

One must ask the question: "Is there any reason to believe that the world behaves according to this class of restricted ARIMA models given in (2.4.1)?" Pandit (1973) has tried to find some theoretical reasons why business, economic, and quality control systems can be predicted by exponential smoothing methods, giving it a "spring-dashpoint" interpretation. The analogy seems strange and is contradicted by many time series, which have been modelled by the three stage iterative Box-Jenkins method.

The data themselves should determine the form of the model and the value of its parameters. The π-weights should depend on the underlying process which has to be identified properly. The automatic and programmed features of exponential smoothing procedures encourage users to

forget that they imply a particular underlying stochastic process. Canned computer programs cannot act as critics, but only as sponsors of the methods they use. It is the responsibility of the individual user to question their validity.

Computation of the forecasts:

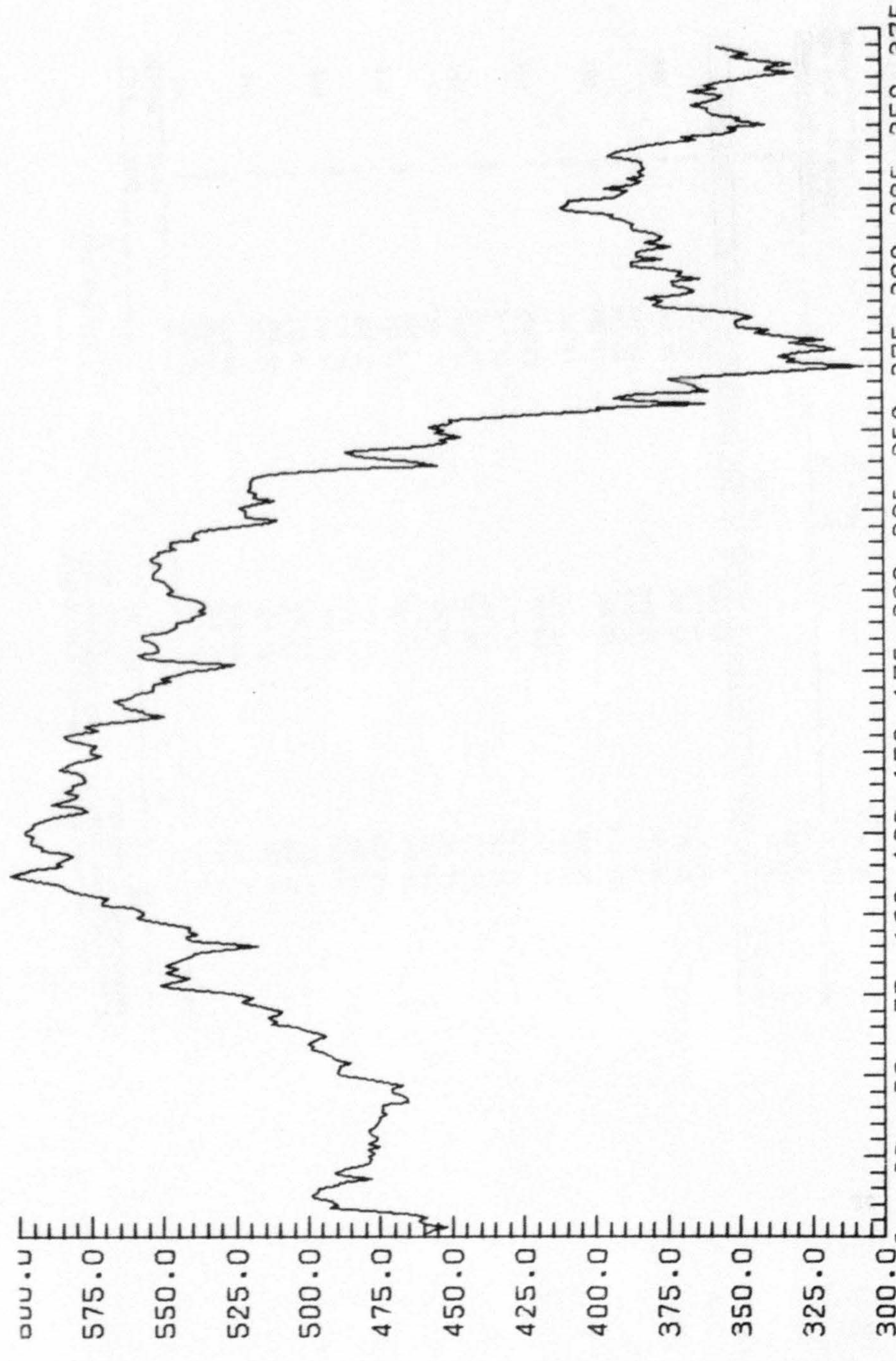
Brown's forecasting procedures are claimed to be computationally efficient. It is easily seen, however, that the forecasts can be derived more readily directly from the difference equation of the equivalent ARIMA model (2.4.1). Thus, even if one believed in the adequacy of Brown's implied model, one should not use his method to calculate and update the forecasts.

These points are best brought out by consideration of specific examples.

2.5 Examples

Example 2.1: Daily IBM common stock closing prices

The data are given on page 526 of Box-Jenkins (1970). In Figure 2.1 the series of 369 observations is plotted. After inspection of the series, Brown (1962) argued that short pieces of the data could be represented by quadratic curves and that one, therefore, ought to consider quadratic fitting functions given by



$$\hat{p}(t,\ell) = b_0(t)+b_1(t)\ell+b_2(t)\ell^2. \qquad (2.5.1)$$

He updates the coefficients of the forecast function by discounted least squares with a smoothing constant of $\beta = .9$.

The forecasts for this form of Brown's model are shown for several time origins and for lead times $\ell = 1,2,3$ in the first column of Table 2.1.

If Brown's model were to be used then it would be much more convenient to use the theory developed in this chapter and to calculate the forecasts directly from the equivalent difference equation given below

$$(1 - B)^3 z_t = (1 - .9B)^3 a_t. \qquad (2.5.2)$$

The forecasts are given in the second column of Tabl 2.1.

In fact, however, as was shown originally by Box and Jenkins, Brown's model seems to be totally inadequate. This is seen for example by the much larger mean squared error of the one, two, and three steps ahead forecasts given in Table 2.1.

Identification using Box and Jenkins methods leads to consider an ARIMA (0,1,1) model with the moving average parameter estimated close to zero

$$(1 - B)z_t = (1 + .087B)a_t. \qquad (2.5.3)$$

time origin	lead time	Brown's model $\hat{p}(t,\ell)=b_0(t)+b_1(t)\ell+b(t)\ell^2$ $\beta = .9$ (2.5.1)	Difference equation form equivalent to Brown's model (2.5.2)	Difference equation form of the ARIMA (0,1,1) suggested by Box-Jenkins (2.5.3)
		F O R	E C A	S T S
300	1	382.59	382.50	376.63
	2	385.56	385.56	376.63
	3	388.66	388.66	376.63
310	1	385.93	385.92	375.96
	2	387.49	387.48	375.96
	3	389.09	389.09	375.96
320	1	409.33	409.32	408.82
	2	411.98	411.98	408.82
	3	414.71	414.70	408.82
330	1	387.31	387.31	384.22
	2	386.72	386.71	384.22
	3	386.07	386.07	384.22
340	1	375.63	375.62	362.92
	2	374.16	374.15	362.92
	3	372.63	372.62	362.92
350	1	346.60	346.60	359.99
	2	344.15	344.14	359.99
	3	341.63	341.62	359.99
360	1	348.21	348.21	342.62
	2	347.06	347.05	342.62
	3	345.89	345.89	342.62
369	1	345.88	345.88	357.38
	2	345.65	345.65	357.38
	3	345.45	345.44	357.38
Lead time		Observed mean squared error of forecasts obtained at various lead times ℓ		
	1		186.81	92.00
	2		279.78	188.00
	3		378.17	291.34
Correlation among the one step ahead forecast errors			$r_1 = .71$, $r_2 = .44$	$r_1 = .00$, $r_2 = .00$

Table 2.1: Forecasts, observed mean squared error of the forecasts and autocorrelation among one step ahead forecast errors for

It was noted that the model in (2.5.3) is very nearly a random walk as originally suggested by Bachelier (1900). This model implies that the best forecasts of future observations are very nearly the current value of the stock price. This is very different from Brown's model which implies that information about the next value is not only contained in the current observation but also in the observations before.

The π-weights for the models (2.5.2) and (2.5.3) are shown in the diagram in Table 2.2.

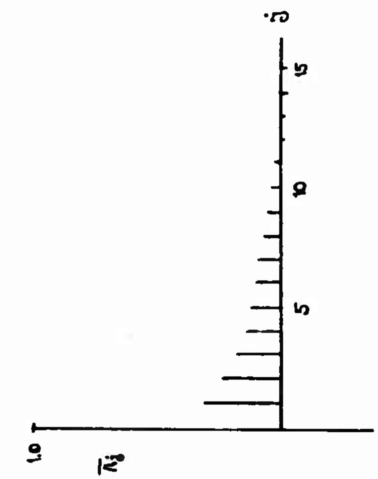
The autocorrelation of the one step ahead forecast errors for models (2.5.2) and (2.5.3) are given in Table 2.1. It can be seen that there is significant autocorrelation among the one step ahead forecast errors for the difference equation model (2.5.2) which is implied by Brown's forecasting procedure in (2.5.1). For the model (2.5.3) the autocorrelations are essentially zero.

Calculating the forecasts:

It is worth emphasizing that if forecasts were to be derived from Brown's model, one should not use Brown's method of calculating them, which is extremely laborious. It is much easier to calculate the forecasts directly

π-weights implied by Brown's model

j	π_j	j	π_j
1	.300	11	.008
2	.240	12	.000
3	.190	13	-.001
4	.146	14	-.001
5	.114	15	-.012
6	.086	16	-.015
7	.063	17	-.016
8	.044	18	-.016
9	.030	19	-.017
10	.018	20	-.016



π-weights of the fitted ARIMA (0,1,1) model (2.5.3) suggested by Box and Jenkins

j	π_j
1	1.087
2	-.094
3	.008
4	-.001
$j \geq 5$	.000

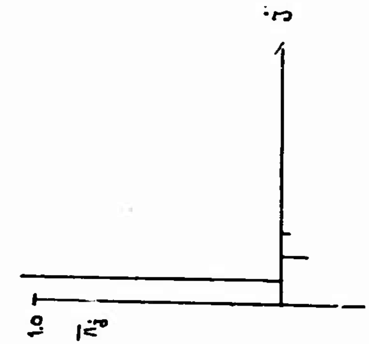


Table 2.2: The π-weights implied by Brown's model (2.5.2) and for the fitted ARIMA (0,1,1) model (2.5.3) suggested by Box and Jenkins.

from the equivalent difference equation. This will give the same result, except for rounding errors, as it is shown in Table 2.1.

The same point can be made in terms of a further example given by Brown.

Example 2.2: Warmdot filter sales

This series is given on page 434 of Brown (1962). In Figure 2.2 the series of 120 observations is plotted.

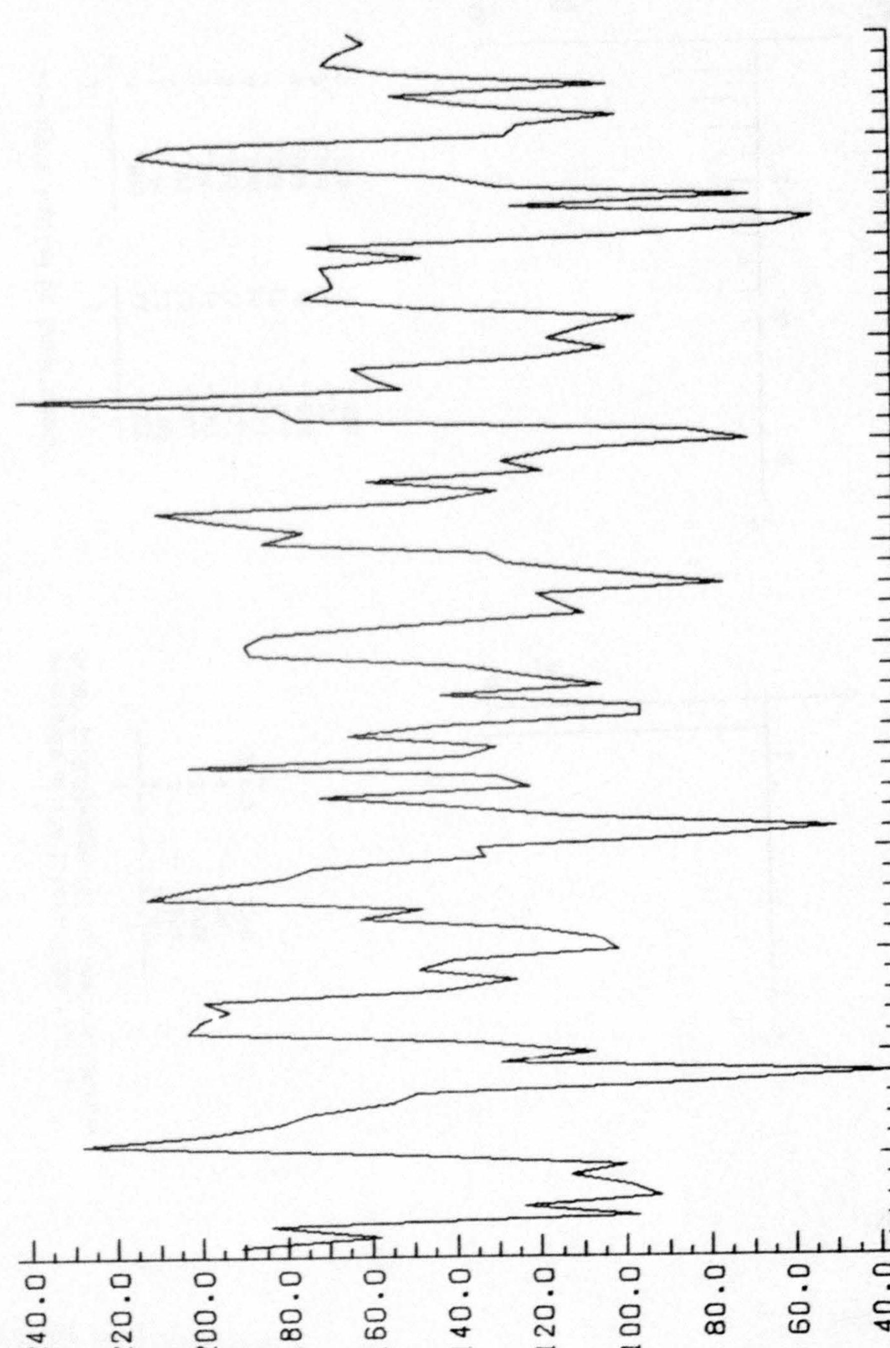
Brown considers the simple 12-point sinusoidal model

$$\hat{p}(t,\ell) = b_0(t)+b_1(t)\sin\frac{2\pi}{12}\ell+b_2(t)\cos\frac{2\pi}{12}\ell \qquad (2.5.4)$$

and updates the coefficients of the forecast function by discounted least squares with smoothing constant $\beta = .9$. The forecasts for this form of Brown's model are shown in the first column of Table 2.3.

There is no point in going through the Box-Jenkins three stage iterative method for this series because this appears to be an artificial series which has been manufactured from the model (2.5.4).

However, it is of interest to see how the theory of this chapter applies to the method of calculating the forecasts. The equivalent difference equation is given by



time origin	lead time	Forecasts for the model (2.5.4) using Brown's method of forecasting with smoothing constant $\beta = .9$	Forecasts for the model (2.5.4) using the equivalent difference equation (2.5.5) for the calculation of the forecasts
80	1	130.01	130.12
	2	151.16	151.27
	3	169.64	169.73
90	1	101.48	101.52
	2	114.61	114.67
	3	135.23	135.28
100	1	105.71	105.72
	2	98.66	98.68
	3	102.36	102.37
110	1	132.17	132.17
	2	102.98	102.98
	3	81.52	81.52
120	1	163.80	163.80
	2	149.53	149.53
	3	133.03	133.03

Table 2.3: Forecasts for the Warmdot filter sales using Brown's model (2.5.4). Comparison of Brown's method of calculating the forecasts and the forecasts derived from the equivalent difference equation (2.5.5).

$$(1-B)(1-\sqrt{3}B+B^2)z_t = (1-.9B)(1-\sqrt{3}(.9)B+(.9)^2B^2)a_t \quad (2.5.5)$$

and forecasts using this difference equation give the same result, except for rounding errors, as shown in the second column of Table 2.3.

The π-weights for the equivalent difference equation model are given in Table 2.4 and indicate how previous observations are used for the calculation of the forecasts.

j	π_j	j	π_j
1	.273	11	.118
2	.153	12	.105
3	.042	13	.077
4	-.036	14	.043
5	-.070	15	.012
6	-.061	16	-.010
7	-.022	17	-.020
8	.030	18	-.017
9	.077	19	-.006
10	.109	20	.008

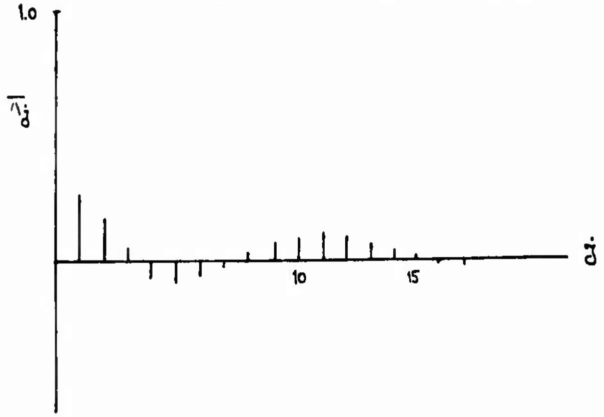


Table 2.4: The π-weights implied by Brown's model (2.5.5); $(1-B)(1-\sqrt{3}B+B^2)z_t=(1-.9B)(1-.9\sqrt{3}B+.81B^2)a_t$.

APPENDIX 2.1

In this appendix we show the following lemma;

<u>Lemma 2.1</u>:

If

$$\underset{[n\times n]}{A} \doteq \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & & a_n \\ \vdots & & & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix} \text{ with } a_i \neq a_j \text{ (for } i \neq j)$$

then

$A^{-1} = \{a_{ij}\}$ where the a_{ij} are given by the expansion in

$$P_i(x) = \prod_{\substack{k=1 \\ k\neq i}}^{n} \frac{x-a_k}{a_i-a_k} = \sum_{j=1}^{n} a_{ij} x^{j-1} .$$

<u>Proof</u>: The existence of the inverse of the Vandermonde matrix is shown by Hoffman and Kunze (1961). We will show that $A^{-1}A = I$.

$$A^{-1}A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_n \\ \vdots & & \vdots \\ a_1^{n-1} & \cdots & a_n^{n-1} \end{bmatrix} = \{v_{ij}\}$$

where $$v_{ij} = a_{i1} + a_j a_{i2} + \ldots + a_j^{n-1} a_{in} . \quad \text{(A.2.1)}$$

v_{ij} is the coefficient of x^0 in the polynomial

$$P_i(x)[1+a_j x^{-1}+a_j^2 x^{-2}+\ldots+a_j^{n-1}x^{-(n-1)}] \quad \text{(A.2.2)}$$

since

$$P_i(x)[1+a_j x^{-1}+a_j^2 x^{-2}+\ldots+a_j^{n-1}x^{-(n-1)}]$$

$$= [a_{i1}x^0+a_{i2}x^1+\ldots+a_{in}x^{n-1}][1+a_j x^{-1}+a_j^2 x^{-2}+\ldots+a_j^{n-1}x^{-(n-1)}]$$

$$= (a_{i1}+a_j a_{i2}+a_j^2 a_{i3}+\ldots+a_j^{n-1}a_{in})x^0$$

$$+ \text{ other terms with } x^\ell \ (\ell \neq 0).$$

Furthermore,

$$P_i(x)[1+a_j x^{-1}+a_j^2 x^{-2}+\ldots+a_j^{n-1}x^{-(n-1)}]$$

$$= \frac{\prod_{\substack{k=1 \\ k\neq i}}^{n} (x-a_k)[1+a_j x^{-1}+\ldots+a_j^{n-1}x^{-(n-1)}]}{\prod_{\substack{k=1 \\ k\neq i}}^{n} (a_i - a_k)} \quad \text{(A.2.3)}$$

$$= \frac{1}{a_i^{n-1} \prod_{k\neq i} (1-\frac{a_k}{a_i})} [x^{n-1}+c_1^{(i)}x^{n-2}+\ldots+c_{n-1}^{(i)}]$$

$$\times[1+a_j x^{-1}+\ldots+a_j^{n-1}x^{-(n-1)}],$$

ere $c_1^{(i)},\dots,c_{n-1}^{(i)}$ are the coefficients in the expansion

$$\prod_{k\neq i}(x-a_k) = x^{n-1}+c_1^{(i)}x^{n-2}+\dots+c_{n-1}^{(i)} . \qquad (A.2.4)$$

e coefficient of x^0 in (A.2.3) is given by

$$v_{ij} = \frac{a_j^{n-1}+a_j^{n-2}c_1^{(i)}+\dots+a_jc_{n-2}^{(i)}+c_{n-1}^{(i)}}{a_i^{n-1}\prod_{k\neq i}(1-\frac{a_k}{a_i})} \div \frac{u_i(j)}{s_i} . \qquad (A.2.5)$$

rthermore,

$$s_i = a_i^{n-1}\prod_{k\neq i}(1-\frac{a_k}{a_i}) = a_i^{n-1}\prod_{k\neq i}(x-\frac{a_k}{a_i})\Big|_{x=1} \text{ evaluated at } x=1$$

t since

$$a_i^{n-1}\prod_{k\neq i}(x-\frac{a_k}{a_i}) = a_i^{n-1}[x^{n-1}+e_1^{(i)}x^{n-1}+e_2^{(i)}x^{n-2}+\dots+e_{n-1}^{(i)}] \qquad (A.2.6)$$

ith $c_\ell^{(i)} = a_i^\ell e_\ell^{(i)}$ for $1 \leq \ell \leq n-1$,

e get from (A.2.6)

$$s_i = a_i^{n-1}[1+\frac{c_1^{(i)}}{a_i}+\frac{c_2^{(i)}}{a_i^2}+\dots+\frac{c_{n-1}^{(i)}}{a_i^{n-1}}] \qquad (A.2.7)$$

$$= a_i^{n-1}+a_i^{n-2}c_1^{(i)}+\dots+a_ic_{n-2}^{(i)}+c_{n-1}^{(i)} .$$

For $i = j$, substituting (A.2.7) into (A.2.5), we get that

$$v_{ii} = \frac{u_i(i)}{s_i} = \frac{a_i^{n-1}+a_i^{n-2}c_1^{(i)}+\dots+a_ic_{n-2}^{(i)}+c_{n-1}^{(i)}}{s_i} = 1 \qquad 1 \leq i \leq n \qquad (A.2.8)$$

For $i \neq j$:

$$\prod_{\substack{k=1\\ k\neq i\ k\neq j\\ i\neq j}}^{n}(x-a_k) = x^{n-2}+f_1^{(i,j)}x^{n-3}+\dots+f_{n-3}^{(i,j)}x+f_{n-2}^{(i,j)} . \qquad (A.2.9)$$

From (A.2.4) it can be easily shown that

$$\left\{\begin{array}{lll} c_1^{(i)} & = f_1^{(i,j)} & - a_j \\ c_2^{(i)} & = f_2^{(i,j)} & - a_jf_1^{(i,j)} \\ \vdots & & \\ c_{n-2}^{(i)} & = f_{n-2}^{(i,j)} & - a_jf_{n-3}^{(i,j)} \\ c_{n-1}^{(i)} & = & - a_jf_{n-2}^{(i,j)} \end{array}\right. \qquad (A.2.10)$$

Thus, for $i \neq j$

$$
\begin{aligned}
u_i(j) &= a_j^{n-1}+a_j^{n-2}c_1^{(i)}+\ldots+a_jc_{n-2}^{(i)}+c_{n-1}^{(i)} \\
&= a_j^{n-1}+a_j^{n-2}[f_1^{(i,j)}-a_j]+a_j^{n-3}[f_2^{(i,j)}-a_jf_1^{(i,j)}] \\
&\quad +\ldots+a_j[f_{n-2}^{(i,j)}-a_jf_{n-3}^{(i,j)}] - a_jf_{n-2}^{(i,j)} = 0
\end{aligned}
$$

and

$$v_{ij} = \frac{u_i(j)}{s_i} = 0 \quad \text{for} \quad i \neq j.$$

q.e.d.

Abraham, B. (1975), "Linear models, time series and outliers," Ph.D. thesis, University of Wisconsin-Madison.

Bachelier, L. (1900), "Theory of Speculation," reprinted in Cootner, P. H. (ed.) [1964], The Random Character of Stock Market Prices," MIT Press, Massachusetts.

Box, G. E. P. and Jenkins, G. M. (1970), Time Series Analysis, Forecasting and Control, Holden-Day; San Francisco.

Brown, R. G. (1962), Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice-Hall; New Jersey.

Dobbie, J. M. (1963), "Forecasting periodic trends by exponential smoothing," Operations Res., 11, 908-918.

Hoffman, K. and Kunze, R. (1961), Linear Algebra, Prentice Hall; Englewood Cliffs, New Jersey.

McKenzie, E. (1973), "A comparison of standard forecasting systems with the Box-Jenkins approach," The Statistician, 23, 107-116.

Pandit, S. M. (1973), "Data dependent systems: modelling, analysis and optimal control via time series," Ph.D. thesis, University of Wisconsin-Madison.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR - TR - 76 - 0722	2. GOVT ACCESSION NO.
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): TOPICS IN TIME SERIES ANALYSIS II. When are exponential smoothing forecast procedures optimal?	5. TYPE OF REPORT & PERIOD COVERED: Scientific Interim
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Johannes Ledolter and George E. P. Box	8. CONTRACT OR GRANT NUMBER(s): AFOSR AF-72-2363-C
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Dept. of Statistics, 1210 W. Dayton Street, Madison, Wisconsin 53706	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research, Building 410, Bolling AFB, D. C. 20332	12. REPORT DATE: Dec., 1975
	13. NUMBER OF PAGES: 41
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

time series; π-weights; ARIMA models; calculation of forecasts; exponential smoothing forecast procedures; eventual forecast function

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper shows that exponential smoothing forecast procedures, in particular those recommended by Brown, will provide optimal (MSE) forecasts only if the underlying process is a member of a particular restricted class of ARIMA models. Actual study of time series, however, does not give any empirical support to this restricted class of models.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE | READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 18 19 AFOSR - TR - 76 - 2722

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle) 6 TOPICS IN TIME SERIES ANALYSIS. II. WHEN ARE EXPONENTIAL SMOOTHING FORECAST PROCEDURES OPTIMAL?

5. TYPE OF REPORT & PERIOD COVERED 9 Interim rept.

6. PERFORMING ORG. REPORT NUMBER Tech Report No 447

7. AUTHOR(s) 10 Johannes Ledolter and George E. P. Box

8. CONTRACT OR GRANT NUMBER(s) 15 AF-AFOSR 72-2363-72

9. PERFORMING ORGANIZATION NAME AND ADDRESS
University of Wisconsin
Department of Statistics
Madison, WI 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
61102F; 9769-05

11. CONTROLLING OFFICE NAME AND ADDRESS
Air Force Office of Scientific Research/NM
Bolling AFB, Washington, DC 20332

12. REPORT DATE 11 December 1975

13. NUMBER OF PAGES 41

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
14 UWIS-DS-75-447

15. SECURITY CLASS. (of this report) 12 23p.
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

16 AF-9769 17 976905

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
time series
ARIMA models
exponential smoothing forecast procedures
eventual forecast function
π-weights
calculation of forecasts

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper shows that exponential smoothing forecast procedures, in particular those recommended by Brown, will provide optimal (MMSE) forecasts only if the underlying process is a member of a particular restricted class of ARIMA models. Actual study of time series, however, does not give any empirical support to this restricted class of models.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED